

Genome-Wide Association Studies of Image Traits Reveal Genetic Architecture of Drought Resistance in Rice

Zilong Guo¹, Wanneng Yang^{1,2,3,*}, Yu Chang¹, Xiaosong Ma⁴, Haifu Tu¹, Fang Xiong¹, Ni Jiang⁵, Hui Feng², Chenglong Huang³, Peng Yang³, Hu Zhao¹, Guoxing Chen⁶, Hongyan Liu⁴, Lijun Luo⁴, Honghong Hu¹, Qian Liu⁵ and Lizhong Xiong^{1,*}

¹National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China

²Hubei Key Laboratory of Agricultural Bioinformatics, Huazhong Agricultural University, Wuhan 430070, China

³College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

⁴Shanghai Agrobiological Gene Center, Shanghai 201106, China

⁵Britton Chance Center for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China

⁶MOA Key Laboratory of Crop Ecophysiology and Farming System in the Middle Reaches of the Yangtze River, Huazhong Agricultural University, Wuhan 430070, China

*Correspondence: Wanneng Yang (ywn@mail.hzau.edu.cn), Lizhong Xiong (lizhongx@mail.hzau.edu.cn)

<https://doi.org/10.1016/j.molp.2018.03.018>

ABSTRACT

Understanding how plants respond to drought can benefit drought resistance (DR) breeding. Using a non-destructive phenotyping facility, 51 image-based traits (i-traits) for 507 rice accessions were extracted. These i-traits can be used to monitor drought responses and evaluate DR. High heritability and large variation of these traits was observed under drought stress in the natural population. A genome-wide association study (GWAS) of i-traits and traditional DR traits identified 470 association loci, some containing known DR-related genes. Of these 470 loci, 443 loci (94%) were identified using i-traits, 437 loci (93%) co-localized with previously reported DR-related quantitative trait loci, and 313 loci (66.6%) were reproducibly identified by GWAS in different years. Association networks, established based on GWAS results, revealed hub i-traits and hub loci. This demonstrates the feasibility and necessity of dissecting the complex DR trait into heritable and simple i-traits. As proof of principle, we illustrated the power of this integrated approach to identify previously unreported DR-related genes. *OsPPI5* was associated with a hub i-trait, and its role in DR was confirmed by genetic transformation experiments. Furthermore, i-traits can be used for DR linkage analyses, and 69 i-trait locus associations were identified by both GWAS and linkage analysis of a recombinant inbred line population. Finally, we confirmed the relevance of i-traits to DR in the field. Our study provides a promising novel approach for the genetic dissection and discovery of causal genes for DR.

Key words: image-based trait, natural variation, GWAS, drought response, drought resistance, rice

Guo Z., Yang W., Chang Y., Ma X., Tu H., Xiong F., Jiang N., Feng H., Huang C., Yang P., Zhao H., Chen G., Liu H., Luo L., Hu H., Liu Q., and Xiong L. (2018). Genome-Wide Association Studies of Image Traits Reveal Genetic Architecture of Drought Resistance in Rice. *Mol. Plant.* **11**, 789–805.

INTRODUCTION

An average annual increase in food production of 44 million metric tons is needed to meet the global food requirement by 2050 (Tester and Langridge, 2010). It is estimated that 1.8 billion people will be confronted with water scarcity by 2025 (Eliasson, 2015). Scarce and unpredictable water resources have made the food shortage situation worse, and enhancing

the drought resistance (DR) of crops is an effective and predictable approach to ensure food security (Hu and Xiong, 2014). Rice (*Oryza sativa* L.) is a staple food crop feeding more than half of the world's population, and is also an important

Molecular Plant

model plant for cereal crops (Xing and Zhang, 2010). Rice production is constrained by drought stress in many rice-producing areas of China, and DR is thus an important trait to improve when developing sustainable rice varieties such as “Green Super Rice” (Zhang, 2007).

However, DR is a complex trait that involves various physiological and molecular responses and can be influenced by a large number of alleles with small effects (Blum, 2011; Fukao and Xiong, 2013). To understand the genetic basis of DR in rice, quantitative trait locus (QTL) mapping using recombinant inbred line (RIL) populations has been conducted (e.g., Yue et al., 2005, 2006). Yield-related traits and visual scores of plant performance during or after drought stress are commonly used to evaluate DR but, remarkably, very few QTLs have been repeatedly detected in different populations or in different years or environments, even for the same population. Grain yield is severely reduced under drought conditions. However, grain yield under drought seems to be controlled by numerous genes with small effect and is affected by many uncontrolled environmental factors in the field. Therefore, it is difficult to reliably identify DR-related loci using grain yield. Other measurements such as manually assessed leaf-drying or leaf-rolling scores tend to be subjective and cannot precisely quantify the level of DR, particularly in large trials.

To enhance reproducibility of DR evaluation in maize, the survival rate after drought stress at the seedling stage was assessed, and this led to the characterization of a previously unrecognized functional DR gene (Wang et al., 2016). However, for many crops, yield is largely determined during the reproductive phase when plants are most sensitive to DR (Alam et al., 2014). A mechanistic understanding of DR mechanisms operative during the reproductive stages could provide useful tools for breeders, as well as being interesting biologically. Therefore, there is an urgent need (i) to develop effective drought phenotyping methods at the reproductive stage; (ii) to understand dynamic drought responses; (iii) to identify highly heritable indicative traits to evaluate DR; and, thereby, (iv) to identify the genetic loci and causal genes underlying variation in these indicative traits.

In recent years, mechanized phenotyping platforms using non-destructive image-based techniques have facilitated the repeated measurement of the same population of individuals as they grow (Busemeyer et al., 2013; Yang et al., 2014). Combining image-based phenomic analyses and genome-wide association studies (GWAS) can reveal the genetic architecture of growth dynamics under normal growth conditions (Bac-Molenaar et al., 2015; Campbell et al., 2017). Similarly, image analysis can be used to understand drought responses (Berger et al., 2010; Chen et al., 2014a; Neumann et al., 2015). GWAS is a powerful approach for understanding the genetic basis of important traits in plants due to its high mapping resolution (Huang et al., 2010, 2011; Zhao et al., 2011). However, no studies have combined image and GWAS analyses to dissect the genetic architecture of DR.

In this study, drought stress was applied to a rice panel (507 accessions) with diversity in DR. Pot-grown rice plants at the reproductive stage were phenotyped before drought stress, under drought stress, and during recovery after rewatering.

Genetic Architecture of Drought Resistance in Rice

Image-based traits (hereafter referred to as i-traits) were measured using a high-throughput and non-destructive phenotyping facility. The i-traits were used to monitor dynamic drought responses and quantify DR. We found that i-traits have high heritability under drought stress. To better reflect DR, we mainly focused on ratio traits (the ratio of the trait value under stress to the trait value before stress). We performed GWAS and established a DR-related association network. A total of 470 loci were identified based on the variation in i-traits and in traditional DR-related traits. These loci included some previously known DR-related genes. Ninety-three percent (437/470) of the loci colocalized with previously reported DR-related QTLs, and 67% (313/470) were identified in two different years. As proof of principle, a previously unreported gene located in a hub locus in the DR-related association network was confirmed to be a causal gene through genetic transformation of rice. In addition to GWAS, we performed linkage analyses for DR using the i-traits in a RIL population. Sixty-nine i-trait locus associations were identified by both GWAS and linkage analyses. In addition, we conducted in-field drought phenotyping in two separate years and demonstrated that several i-traits and associated loci could be useful for DR improvement in the field. Finally, we developed an open-access database that includes all the images and genotypic and phenotypic data from this study. This database is aimed at encouraging data-reuse by the plant science and computer vision communities, which will be useful to collectively understand and exploit the abundant natural variation of DR.

RESULTS

Capturing Drought Responses by Using a Phenotyping Platform

To understand the dynamic drought responses of rice plants, we acquired images of the variety Swarna (*Oryza sativa* L. ssp. *indica*) at eight time points during progressive drought stress and rewatering (Figure 1A and 1B), using the rice automatic phenotyping platform (RAP) described by Yang et al. (2014). After image processing, 51 image-based traits (hereafter referred to as “i-traits”) were extracted (Table 1 and Methods). These i-traits were classified into five categories: biomass-related traits, leaf stay-green-related traits, morphological traits, histogram texture traits, and texture traits derived from a gray-level co-occurrence matrix.

We found that the change in i-traits reflected drought progression (Figure 1C). These dynamic changes followed three distinct patterns: (i) drought responses observed only under severe drought conditions (such as GPAR, which is defined as the “green projected area ratio”); (ii) similar changes observed under both moderate and severe drought conditions (such as TPA, defined as the “total projected area”); and (iii) different responses observed under moderate and severe drought conditions (such as GPA, defined as the “green projected area”) (Figure 1D). Dynamic changes in other i-traits are shown in Supplemental Figure 1. These results suggest that the i-traits are useful for the study of dynamic drought responses.

Variation of I-Traits in a Natural Diversity Population

To investigate the variation of i-traits during drought responses, we phenotyped a population of 507 rice accessions, collected a

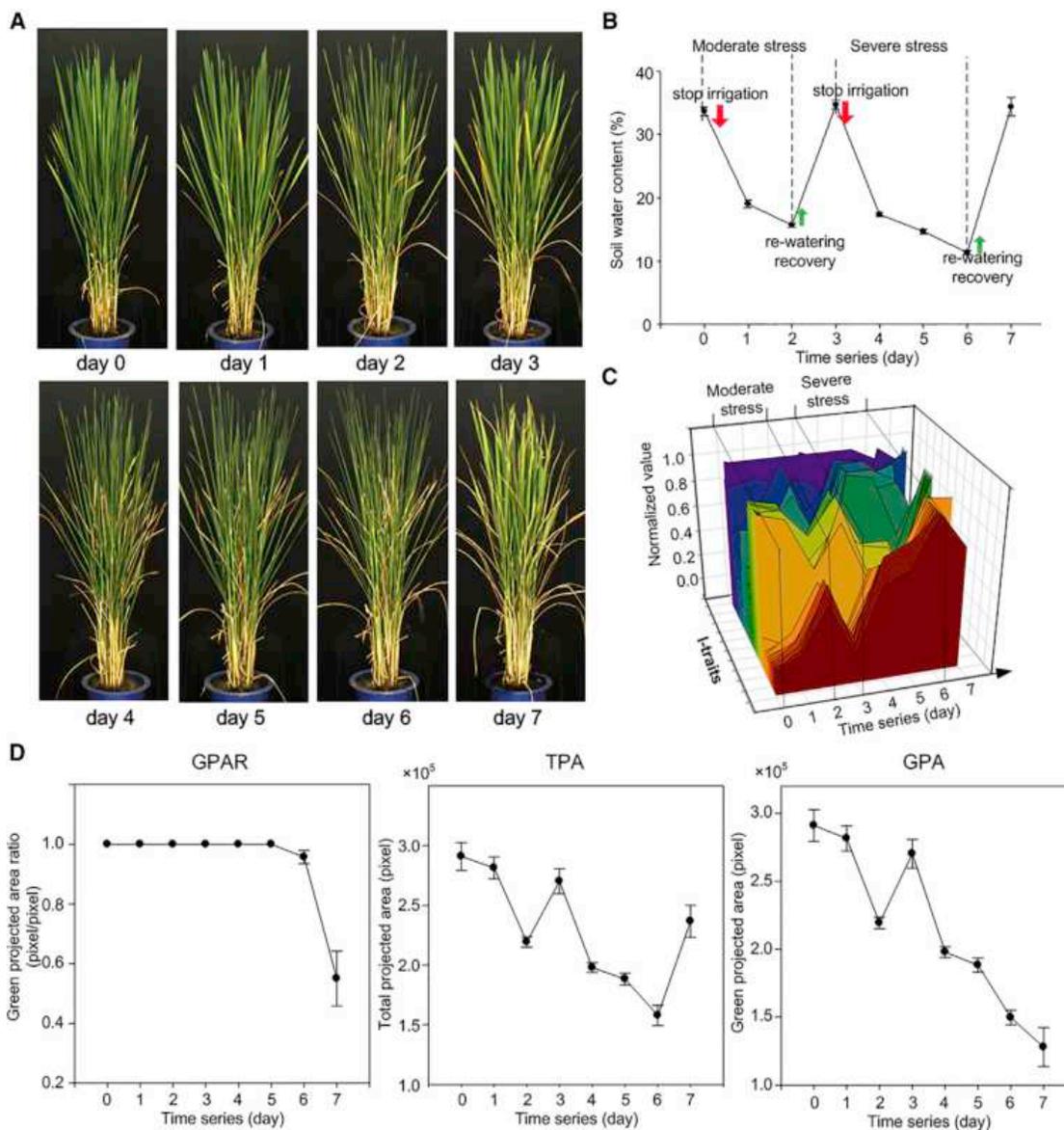


Figure 1. Drought Response Monitored by 51 I-Traits Measured by RAP.

(A) Raw images of a rice variety (Swarna) measured at eight time points during progressive drought stress and rewatering.

(B) Soil water content measured by TDR at eight time points. Two rounds of drought stress were applied to three plants of the rice variety Swarna. Based on soil water content, the number of days of drought stress, and the stay-green level of the stressed plants after rewatering, the levels of stress, i.e. “moderate stress” (2 days for the first round of stress) and “severe stress” (3 days for the second round of stress), were determined. Error bars indicate the SE based on three biological replicates.

(C) Normalized value of 51 i-traits during progressive drought stress and rewatering. The range of values at eight time points for each i-trait was transformed to 0–1 by linear normalization. $y = (x - \min) / (\max - \min)$ where x , y , \max , and \min represent raw data, normalized data, maximum, and minimum, respectively.

(D) Three temporal patterns of i-traits (GPAR, TPA, and GPA) during progressive drought stress and rewatering. GPAR, green projected area ratio; TPA, total projected area; GPA, green projected area. Error bars indicate the SE based on three biological replicates.

total of 70 980 images, and measured 51 i-traits at three distinct stages (before drought stress, represented by the suffix “_B”; after 5 days of drought treatment with a soil water content of about 15%, represented by the suffix “_D”; and after complete recovery, represented by the suffix “_Re”) (Figure 2A, Supplemental Table 1, and Methods). The values of all 51 i-traits under drought stress were significantly different from those before drought stress (P values ranging from 2.27×10^{-3}

to 9.54×10^{-258} , paired-samples t -test) (Figure 2B). To better reflect drought response, we focused on ratio traits (represented by the suffix “_R”) calculated as the ratio of the i-trait value under drought stress to the value before drought stress. Based on i-traits_R for 507 rice accessions, large variation in DR was observed (Figure 2B). For example, GPAR, which is defined as the ratio of the green projected area to the TPA of a plant and was used to quantify the proportion of

Classification	I-trait	Definition	Additional annotation
Biomass-related	TPA	Total projected area	Shoot weight/biomass
Greenness-related	GPA	Green projected area	Reflects stay-green
	GPAR	Green projected area ratio	Reflects stay-green
	LGPA	Light-green projected area	Reflects stay-green
	DGPA	Dark-green projected area	Reflects stay-green
Morphological-related	FDNIC	Fractal dimension without image cropping	
	FDIC	Fractal dimension after image cropping	
	H	Height of the bounding rectangle of the plant	Plant height
	W	Width of the bounding rectangle of the plant	Plant width
	HWR	Height/width ratio	Reflects tiller angle
	TBR	Total projected area/bounding rectangle area ratio	Reflects leaf rolling
	PAR	Perimeter/projected area ratio	Reflects leaf rolling
	PC1–PC6	Plant compactness	PC6 reflects plant compactness
Histogram texture-related	F1–F14	Relative frequencies	
	M	Mean value	
	SE	Standard error	
	MU3	Third moment	
	U	Uniformity	
	S	Smoothness	
Gray-level co-occurrence matrix texture-related	E	Entropy	
	T1	Correlation	
	T2	Advantages of the small gradient	
	T3	Advantages of the large gradient	
	T4	Energy	
	T5	Intensity inhomogeneity	
	T6	Gradient inhomogeneity	Reflects stay-green
	T7	Mean gray	
	T8	Mean gradient	
	T9	Gray entropy	
	T10	Gradient entropy	
	T11	Entropy of mixing	
	T12	Differential moment	
	T13	Deficit score	
	T14	Gray mean variance	
T15	Gradient mean variance		

Table 1. I-Traits Measured by RAP.

healthy non-senescent tissue under drought stress, significantly differed between the three stages. There was no obvious variation in GPAR between accessions before drought stress (the coefficient of variation was 0.09), but GPAR decreased under drought stress and the variation between accessions increased dramatically ($P = 2.38 \times 10^{-254}$, *t*-test; the coefficient of variation was 0.71). After rewatering, GPAR failed to recover but TPA, did recover (Figure 2C). Stay-green, defined as the ability to counteract leaf senescence caused by drought stress (Rosenow et al., 1983), has been commonly used as a measure

of DR and is typically quantified through visual scoring based on the degree of leaf drying (Crasta et al., 1999; Yue et al., 2006). Instead of a visual rating, which can be observer biased, the i-trait GPAR can be used to objectively quantify stay-green. Most of the i-traits (77.8%) showed high repeatability ($w^2 > 0.7$) at all three stages. We developed an open-access phenomic database containing the large amount of data generated in this study, which can be used for integrative analyses of DR with other omics datasets (http://plantphenomics.hzau.edu.cn/search_en.action).

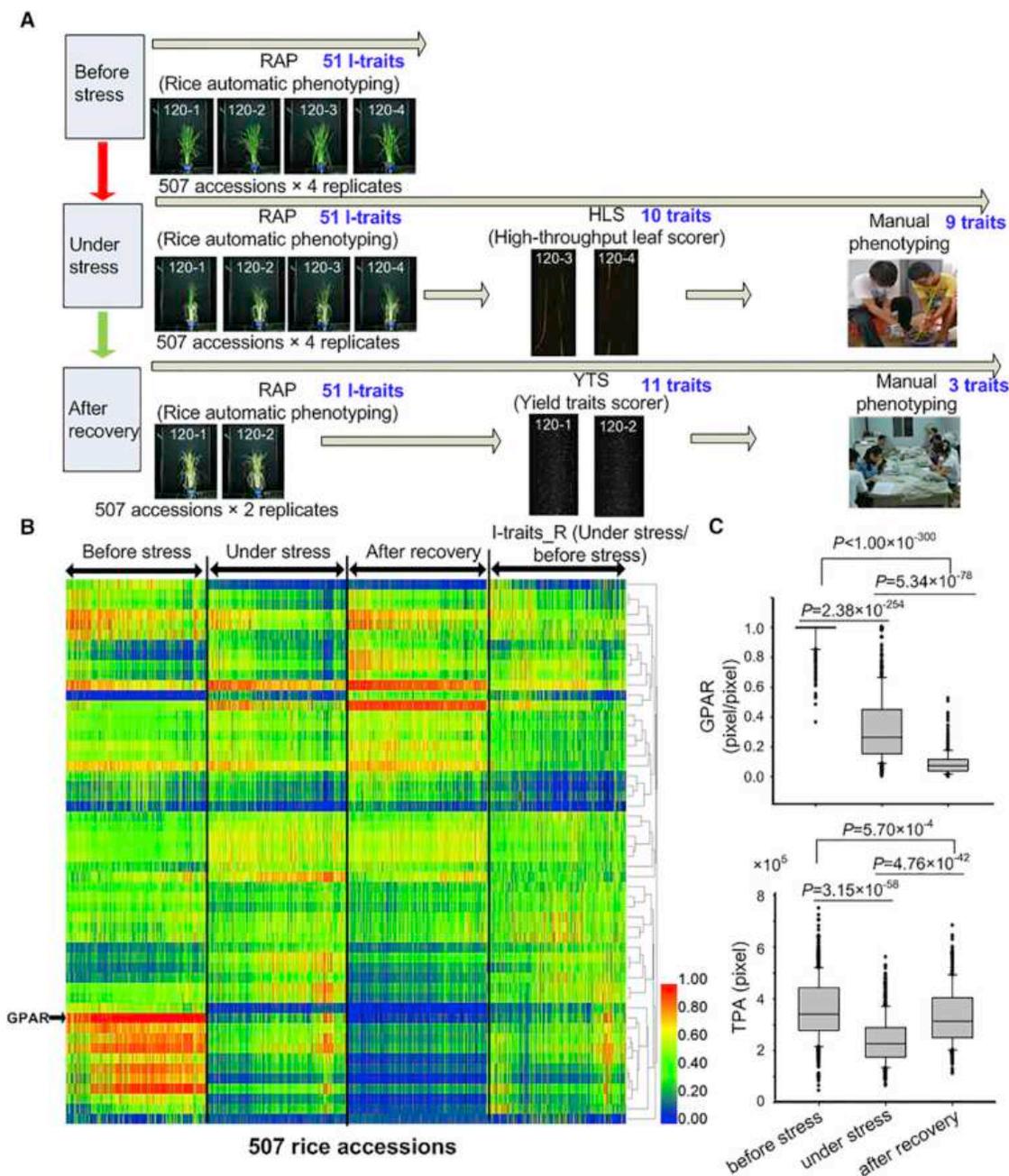


Figure 2. Drought Phenotyping of a Natural Population and Variation of I-Traits.

(A) Flow chart showing the procedure for drought phenotyping of 507 rice accessions. The phenotypic data were collected at three different stages (before drought stress, after 5 days of drought treatment with a soil water content of about 15%, and after complete recovery) using RAP (for i-traits), HLS (for leaf traits), and YTS (for yield-related traits), and manual measurement.

(B) Heatmap showing the values of 507 accessions before drought stress, under drought stress, and after rewatering, and i-traits_R (under drought stress/before drought stress). Before drawing the heatmap, the range of values from 507 accessions for each i-trait was transformed to 0–1 by linear normalization. $y = (x - \min) / (\max - \min)$ where x , y , \max , and \min represent raw data, normalized data, maximum, and minimum, respectively.

(C) Box plot showing the variation of GPAR, TPA, and the difference between these traits at different stages. P values for pairwise comparisons of drought stages were calculated using t -tests.

Relevance of I-Traits to Traditional DR-Related Traits

In addition to the i-traits measured non-destructively on the RAP platform, we measured several traits traditionally used for DR evaluation including leaf stay-green, leaf water content, and grain yield-related traits (Figure 2A and Supplemental Table 1). For stay-green traits, we quantified the greenness of leaves by using

a high-throughput leaf scorer (HLS) described by Yang et al. (2015). We measured grain yield traits in a highly efficient manner by using a yield trait scorer (YTS) described by Yang et al. (2014). To quantify DR in terms of plant productivity, we used eight yield-related traits under normal and drought conditions to calculate relative values (trait values under drought

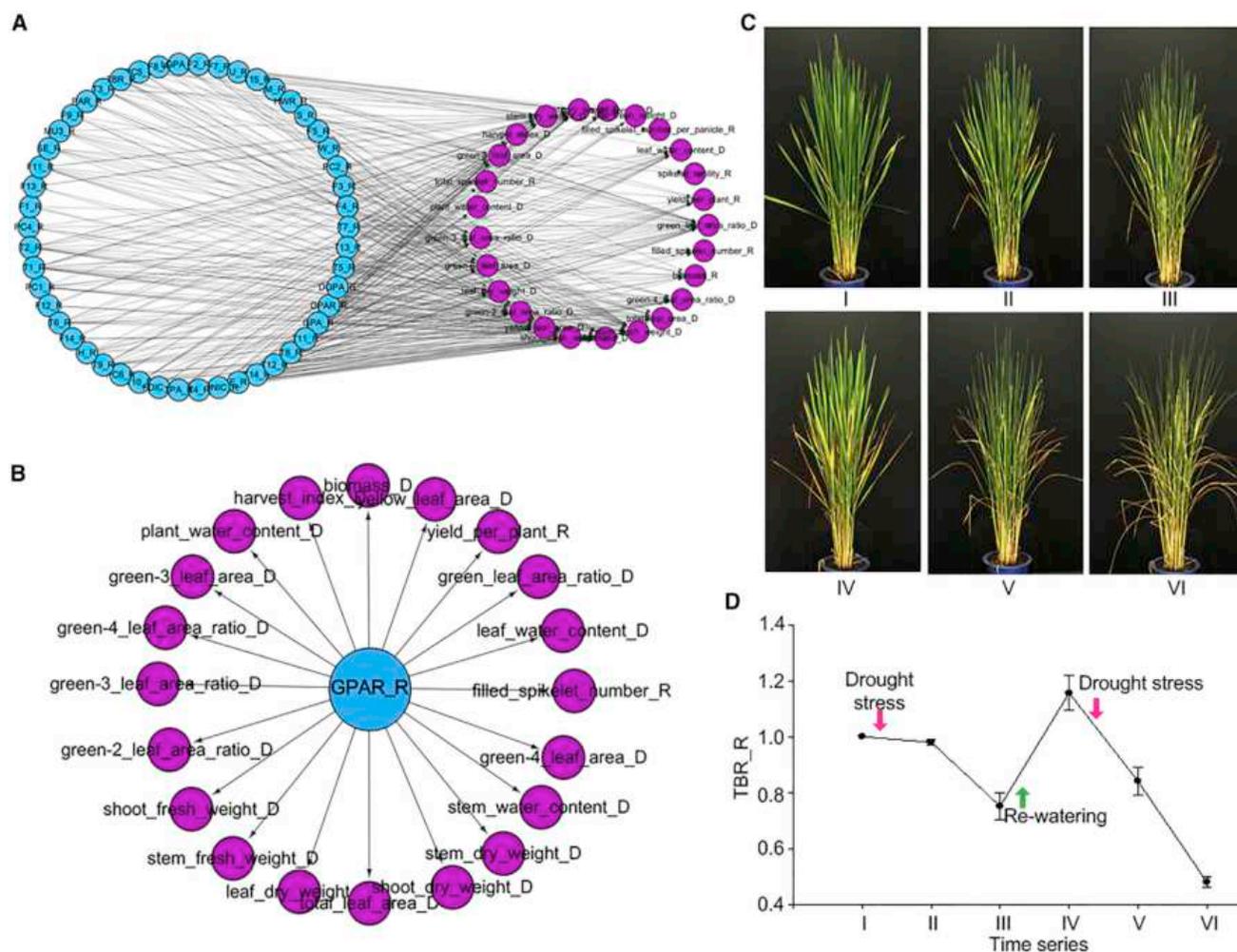


Figure 3. Relevance of I-Traits to Traditional DR-Related Traits.

(A) Correlation network of all ratio i-traits (i-trait values under stress/i-trait values before stress) and traditional DR-related traits. (B) Correlation network of GPAR_R and traditional DR-related traits. Blue circles, amaranth circles, and lines with arrowheads represent i-trait_R, traditional DR-related traits, and the correlations between them (Pearson correlation coefficient $R \geq 0.3$, $P < 0.001$), respectively. (C) Raw images of a rice variety (Swarna) taken at six time points during progressive drought stress and rewatering to show dynamic leaf rolling caused by drought stress. (D) Value of TBR_R at six time points corresponding to images I–VI of (C). TBR_R is the ratio of TBR at a given time point to TBR at the first time point (before stress). Error bars indicate the SE based on three biological replicates.

conditions/trait values under normal conditions, represented by the suffix “_R”) for the following analyses. In the diversity population, large variation of traditional DR-related traits was also observed. For example, , under the same drought conditions (i.e. the same soil water content) the leaf water content in the population ranged from 6.3% to 72.6%, indicating large variation in drought avoidance capability (Supplemental Table 2).

Systematic correlation analyses between the i-traits_R and the traditional DR traits were done to construct a correlation network (phenotypic Pearson’s correlation coefficient $R \geq 0.3$, $P < 0.001$) (Riedelheimer et al., 2012; Xie et al., 2015). This network contained 390 associations between the i-traits_R and the traditional DR traits (Figure 3A and Supplemental Table 3). On average, each i-trait_R was correlated with eight traditional DR traits. For example, GPAR_R had 20 associations with traditional traits, including the green leaf area ratio at the whole-

plant level under drought stress ($R = 0.66$, $P < 0.001$), the water content of leaves under stress ($R = 0.49$, $P < 0.001$), the harvest index under drought stress ($R = 0.43$, $P < 0.001$), the relative yield per plant ($R = 0.36$, $P < 0.001$), and the relative filled spikelet number ($R = 0.33$, $P < 0.001$) (Figure 3B). We found that the change in the i-trait TBR_R (total projected area/bounding rectangle area ratio) was strongly correlated with the degree of leaf rolling under progressive drought stress and rewatering (Figure 3C and 3D). We examined the raw images of 507 rice accessions and found that most accessions with decreased TBR ($TBR_R < 1$) showed leaf rolling under drought stress while those with increased TBR ($TBR_R > 1$) showed leaf wilting without leaf rolling. Furthermore, we found that TBR_R was significantly correlated with the actual leaf-rolling score ($R = -0.84$, $P = 3.39 \times 10^{-29}$). These results indicate that TBR_R can be used to quantify the degree of leaf rolling.

We constructed optimized models to predict the shoot weight under drought stress and after rewatering using all-subset regression. The models were evaluated using 5-fold cross-validation. The adjusted determination coefficients (adjusted R^2) of the models containing only the i-trait TPA were larger than 0.70, indicating that TPA can be used to predict biomass (Supplemental Table 4). We also found that some i-traits collectively could explain the variance of traditional DR traits (percentage of variance explained ranging from 20.2% to 78.8% using linear stepwise regression) (Supplemental Table 5). For example, 17 i-traits in combination explained 54.8% of the variance in the leaf water content under drought stress; 25 i-traits explained 74.5% of the variance in the green leaf area ratio at the whole-plant level under drought stress. These results suggest that the i-traits can be used to not only monitor dynamic drought responses, but also to evaluate overall DR.

Genome-Wide Association Study of DR

We performed a GWAS of 255 traits (where the same trait at different stages was regarded formally as a different trait) using a linear mixed model. After Bonferroni correction, performed as in previous studies, the genome-wide P value thresholds were set to 1.21×10^{-6} , 1.66×10^{-6} , and 3.81×10^{-6} for the whole population, *indica* subpopulation, and *japonica* subpopulation, respectively (Yang et al., 2014, 2015; Wang et al., 2015). In total, 6842 associations between 3379 lead SNPs (the SNP with the lowest P value for a particular association signal) and 252 traits were identified in at least one population. Considering the slow linkage disequilibrium (LD) decay in rice and based on previously published studies (Chen et al., 2014b; Yang et al., 2015; Crowell et al., 2016), adjacent lead SNPs within a region less than 300 kb were defined as a single locus. A total of 617, 403, 270, and 443 loci were detected using i-traits_B (before drought), i-traits_D (drought), i-traits_Re (recovery), and i-traits_R (i-traits_D/i-traits_B), respectively (Figure 4A). We mainly focused on the GWAS results for the ratio traits (represented by the suffix “_R”) because they theoretically better reflect DR than trait values under drought stress (represented by the suffix “_D”). We identified a total of 1074 associations between the 59 ratio traits (including i-traits and traditional DR-related traits) and 793 lead SNPs (corresponding to 470 loci) (Supplemental Table 6). Among the 470 DR-related loci, 94% (443/470) were identified using i-traits_R, and these loci were not randomly distributed in the genome ($\chi^2 = 660.08$, $P = 6.29 \times 10^{-62}$), indicating the existence of hotspot regions where DR-related loci were enriched (Figure 4B).

Previously reported DR-related QTLs were retrieved from the TropGeneDB (Ruiz et al., 2004; Hamelin et al., 2013), QTARO (Yonemaru et al., 2010), and PubMed (see Methods) databases. Ninety-three percent (437/470) of associated loci identified in this study co-localized with previously reported DR-related QTLs, including many QTLs controlling yield traits under drought conditions in the field (Supplemental Table 7). For example, a locus on chromosome 3 was associated with GPAR_R and harvest index_R based on the GWAS results, and this locus co-localized with a reported QTL controlling harvest index under severe drought stress in the field (Lanceras et al., 2004); another locus on chromosome 2 associated with TBR_R co-localized with a reported QTL controlling leaf-rolling

score under drought stress in the field (Price et al., 2002). These results suggest that the DR-related loci identified in our study are reliable and could be useful for DR breeding selection.

Some known DR-related genes were significantly associated with i-traits. For example, *OsWRKY13*, which encodes a transcription factor, negatively regulates DR (Xiao et al., 2013). One SNP in the 1-kb promoter region, one SNP in 5' UTR, and two SNPs in the coding region causing amino acid changes were significantly associated with the i-trait GPAR_R ($P_{LMM} = 3.07 \times 10^{-4}$, 6.76×10^{-5} , 1.08×10^{-4} , and 1.08×10^{-4} , respectively). GPAR_R significantly differed between eight *OsWRKY13* haplotype groups ($P = 2.16 \times 10^{-12}$, Kruskal-Wallis ANOVA). Haplotype H1 was the superior haplotype and was mainly found in *japonica* accessions (Figure 4C). Another example is *OsDREB2A*, which encodes an AP2/EREBP transcription factor that positively regulates DR (Cui et al., 2011). Fifteen SNPs in the promoter region and one SNP in the 5' UTR were significantly associated with GPAR_R ($P_{LMM} < 1 \times 10^{-4}$); GPAR_R significantly differed between eight *OsDREB2A* haplotype groups ($P = 7.99 \times 10^{-8}$, Kruskal-Wallis ANOVA). Both H1 and H2 (mainly found in *japonica* accessions) were the superior haplotypes (Figure 4D). These results further support the usefulness of GPAR_R in the genetic dissection of DR.

DR-Related Association Network Based on GWAS Results

We further constructed a DR-related association network based on GWAS results to identify hub loci and hub traits for DR. We found that each locus was associated with two ratio traits on average (ranging from 1 to 11) and each ratio trait was associated with 16 loci on average (ranging from 1 to 75). We defined loci with ≥ 5 associations and traits with ≥ 20 associations as hub loci and hub traits, respectively. Thus, 37 hub loci and 18 hub traits (including 17 i-traits) were identified. Based on the association network, we found that DR-related yield traits (such as spikelet fertility_R) were involved in at least two association clusters that contained hub traits GPAR_R and TBR_R, respectively (Figure 5A). GPAR_R and TBR_R can be used to quantify stay-green and leaf rolling at the whole-plant level under drought stress, respectively (Figure 5B). Stay-green and leaf rolling largely reflect drought tolerance (DT) and drought avoidance (DA) mechanisms, respectively (Yue et al., 2006; Fang and Xiong, 2015; Johnson et al., 2015). Strong association signals with clear peaks were identified for the two traits (42 loci associated with GPAR_R and 20 loci associated with TBR_R) (Figure 5C), but only one locus overlapping both traits was found. Based on these observations, we propose that GPAR_R and TBR_R largely reflect DT and DA, respectively. To test this hypothesis, we retrieved reported rice QTLs controlling DR-related traits (including leaf-drying score, osmotic adjustment, cell membrane stability, leaf-rolling score, canopy temperature, and leaf relative water content under drought conditions) from the database TropGeneDB. We found that most of the loci associated with GPAR_R and TBR_R overlapped with QTLs controlling DT and DA-related traits, respectively (Supplemental Table 8). For GPAR_R, 66.7% (18/27) of the loci overlapped with previously known QTLs controlling DT-related traits such as leaf-drying score, osmotic adjustment, and cell membrane stability. For TBR_R, 83.3% (10/12) of the loci overlapped with

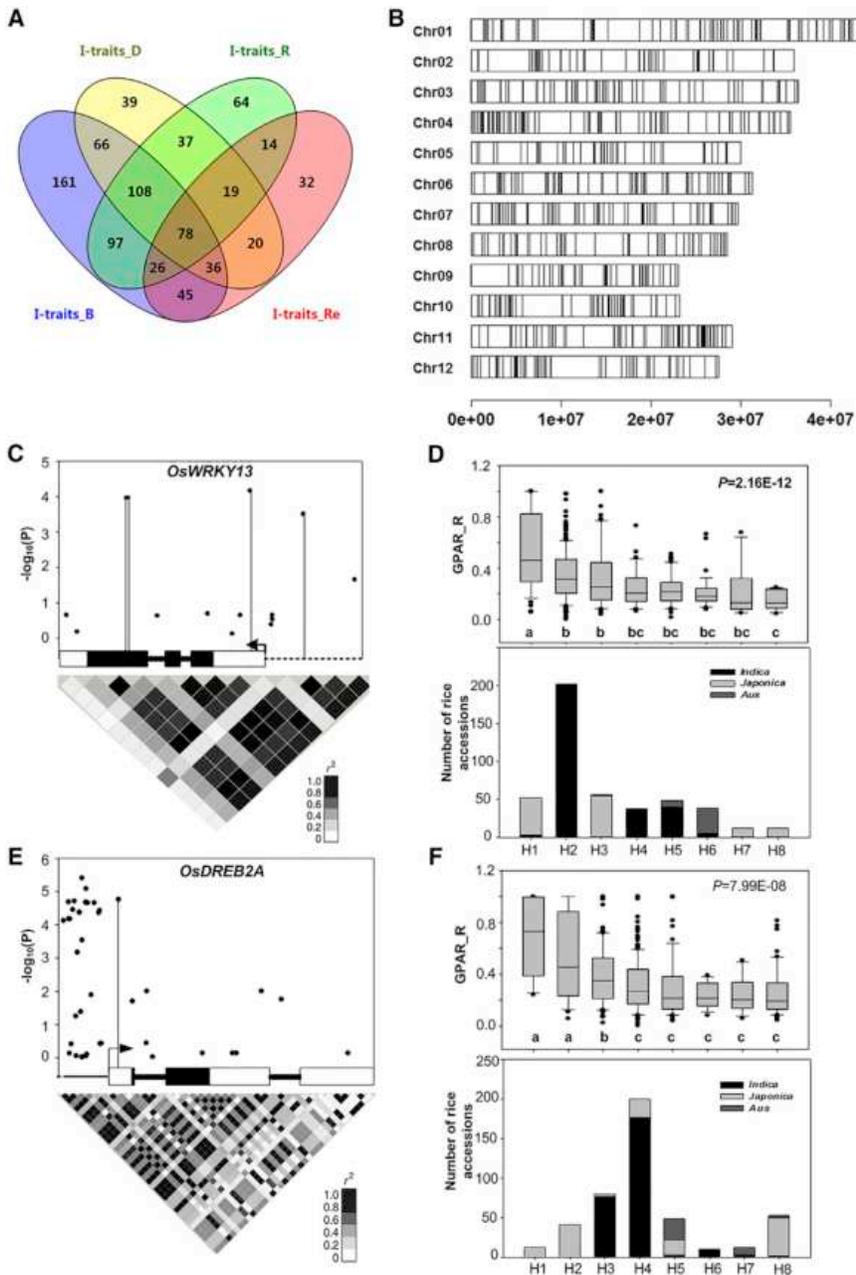


Figure 4. DR-Related GWAS Results.

(A) Venn diagram showing the number of loci associated with i-traits before drought stress (I-traits_B), under drought stress (I-traits_D), after rewatering (I-traits_Re), and ratio i-traits (I-traits_R).

(B) Chromosomal distribution of loci associated with ratio traits (including i-traits and traditional traits). Each vertical line indicates a lead SNP.

(C) Local Manhattan plot and LD statistic r^2 values for a known DR-related gene, *OsWRKY13* (gene body and 1-kb upstream region), associated with GPAR_R.

(D) Comparison of GPAR_R among eight haplotype groups and the number of accessions included in each *OsWRKY13* gene haplotype group.

(E) Local Manhattan plot and LD statistic r^2 values for a known DR-related gene, *OsDREB2A* (gene body and 1-kb upstream region), associated with GPAR_R.

(F) Comparison of GPAR_R among eight haplotype groups and the number of accessions included in each *OsDREB2A* gene haplotype group.

The dotted line, solid lines, white rectangles, and black rectangles represent the 1-kb promoter region, introns, UTRs, and exons, respectively. The arrow indicates the transcription start site and transcription orientation. The haplotypes were determined using the SNPs with a MAF ≥ 0.05 . The distribution of i-traits for each haplotype group is shown in the box plot. The i-trait values of multiple haplotype groups were compared using Kruskal–Wallis one-way ANOVA. The number of rice accessions in each haplotype group is shown on the histogram.

For the Manhattan plot, $-\log_{10} P$ values calculated using a mixed linear model are plotted against the position of the SNPs.

Identification of a New DR-Related Gene

T6 is a texture trait derived from a gray-level co-occurrence matrix. T6_R (ratio of T6 under stress/before stress), which ranked second in the list of hub traits, was positively correlated with the green leaf area ratio under stress (Pearson correlation coefficient

QTLs controlling DA-related traits such as leaf-rolling score, canopy temperature, and leaf relative water content under drought conditions.

Of the 470 loci in the DR-related association network, 443 loci were identified using i-traits and 52 loci were identified using traditional DR traits. The low number of loci identified using traditional DR traits is perhaps due to the low mapping power of these traits. When using a less stringent genome-wide P value threshold (1.00×10^{-3}) for traditional DR traits, 127 out of 443 loci associated with i-traits were also associated with traditional DR traits. For example, the lead SNP sf1016269098 was associated with GPAR_R ($P_{LMM} = 1.61 \times 10^{-6}$) and yield_R ($P_{LMM} = 1.07 \times 10^{-4}$), and the genotype C SNP allele was the superior allele for both GPAR_R and yield_R (Figure 5D).

$R = 0.42$, $P < 0.001$), leaf water content under stress ($R = 0.38$, $P < 0.001$), yield_R ($R = 0.29$, $P < 0.001$), and biomass_R ($R = 0.41$, $P < 0.001$). These results indicate that T6 may also be related to the stay-green trait. Two known DR-related genes, *OsNCED3* (Cai et al., 2015) and *OsDREB1E* (Chen et al., 2008), were located in the loci associated with T6_R. Using RiceNet (rice gene network prioritization web server) (Lee et al., 2015), we found that an unreported gene (*OsPP15*, LOC_Os01g62760) was associated with these two known genes (see Methods). *OsPP15* encodes a PP2C homolog of ABI1, a key member in the abscisic acid (ABA) signaling pathway (Meyer et al., 1994), and was associated with T6_R ($P_{LMM} = 1.63 \times 10^{-4}$ for SNP sf0136344957) (Figure 6A). T6_R significantly differed between seven *OsPP15* haplotype groups ($P = 5.18 \times 10^{-8}$,

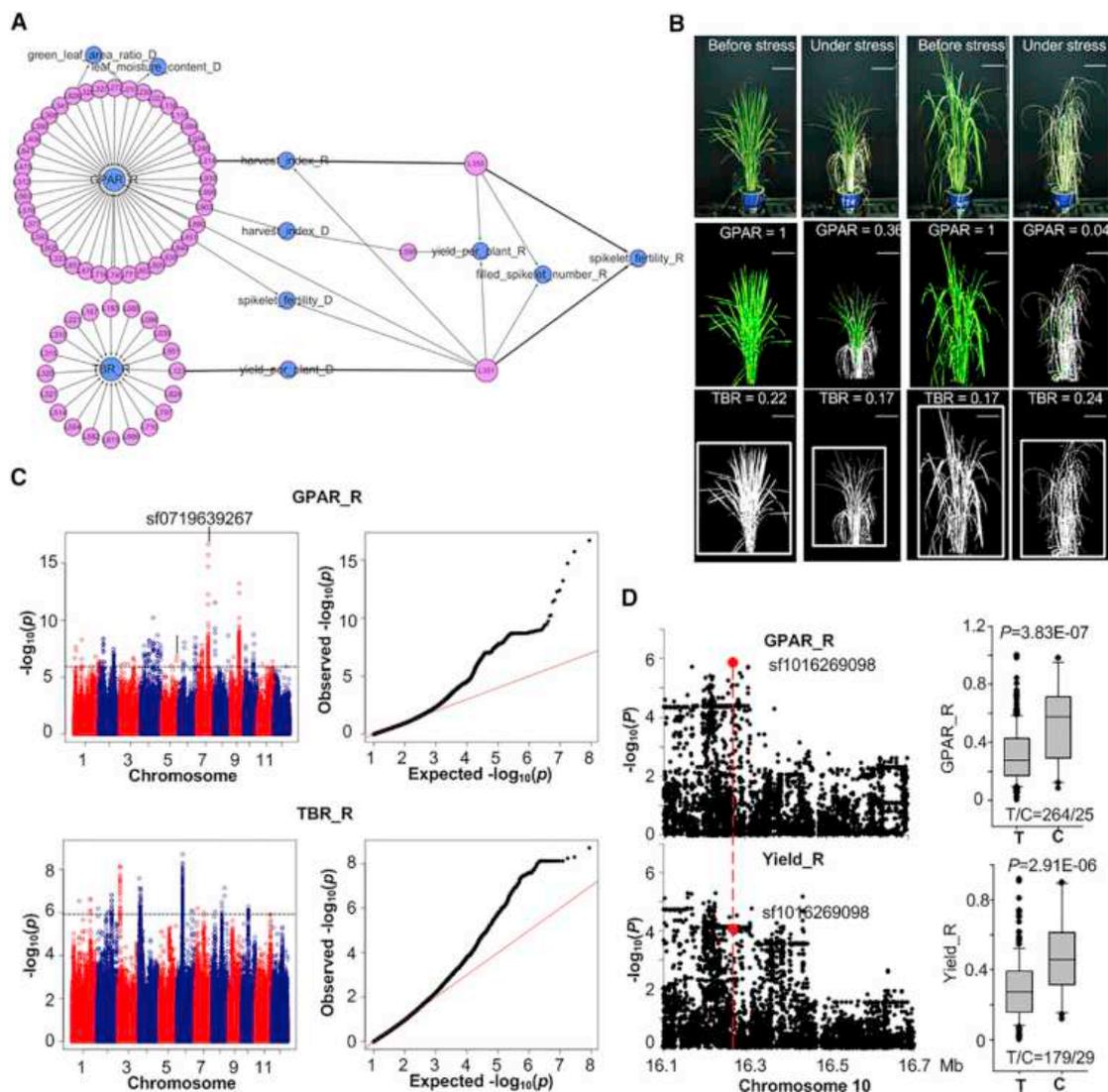


Figure 5. Association Network Based on GWAS Results.

(A) Local association network containing two main association clusters containing the hub traits GPAR_R and TBR_R. The pink and blue circles represent genetic loci and DR-related traits, respectively; the lines represent significant associations between loci and traits.

(B) Raw images of two rice accessions taken before stress and under stress (first row); green plant pixels and GPAR (second row); segmented images and bounding rectangle of a plant and TBR (third row). Scale bars, 0.2 m.

(C) GWAS plots for GPAR_R and TBR_R. “_R” indicates that the trait value is the ratio of stress/non-stress values. For the Manhattan plot (left), $-\log_{10} P$ values from a genome-wide scan are plotted against the position of the SNPs on each of 12 chromosomes, and the horizontal gray dotted line indicates the genome-wide P value threshold; for the quantile–quantile plot (right), the horizontal axis shows the expected $-\log_{10}$ -transformed P values, and the vertical axis indicates the observed $-\log_{10}$ -transformed P values.

(D) Manhattan plots of the same genomic region associated with GPAR_R and Yield_R (left panel). Based on the genotype of SNP sf1016269098, the rice accessions were classified into two groups. The distributions of GPAR_R and Yield_R for the two groups are shown in the box plot. The trait values of the two groups were compared using t -tests.

Kruskal–Wallis ANOVA). Haplotype H1 (primarily found in *indica* accessions), and H7 (primarily found in *japonica* accessions) were superior and inferior haplotypes, respectively (Figure 6B). To confirm *OsPp15* as a causal gene, it was overexpressed in rice Zhonghua 11. Measurements of leaf relative water content, relative electrolyte leakage, and survival rate (Figure 6C and 6D) indicated that the *OsPp15*-overexpression lines were hypersensitive to drought stress, which is in agreement with the predicted role of *OsPp15* as a negative regulator of DR.

Validation by Repeated GWAS in Different Years and Linkage Analyses

To further validate the reliability of *i*-traits in DR studies, we selected 300 diverse accessions from the association population for drought phenotyping, again using the RAP platform, in 2016. Based on the phenotypic data from two years (2013 and 2016), 41.2% (21/51) of *i*-traits_R showed high heritability ($H^2 > 0.5$) while H^2 values for all the traditional DR ratio traits were less than 0.5 (Supplemental Table 9). For example, the heritability of

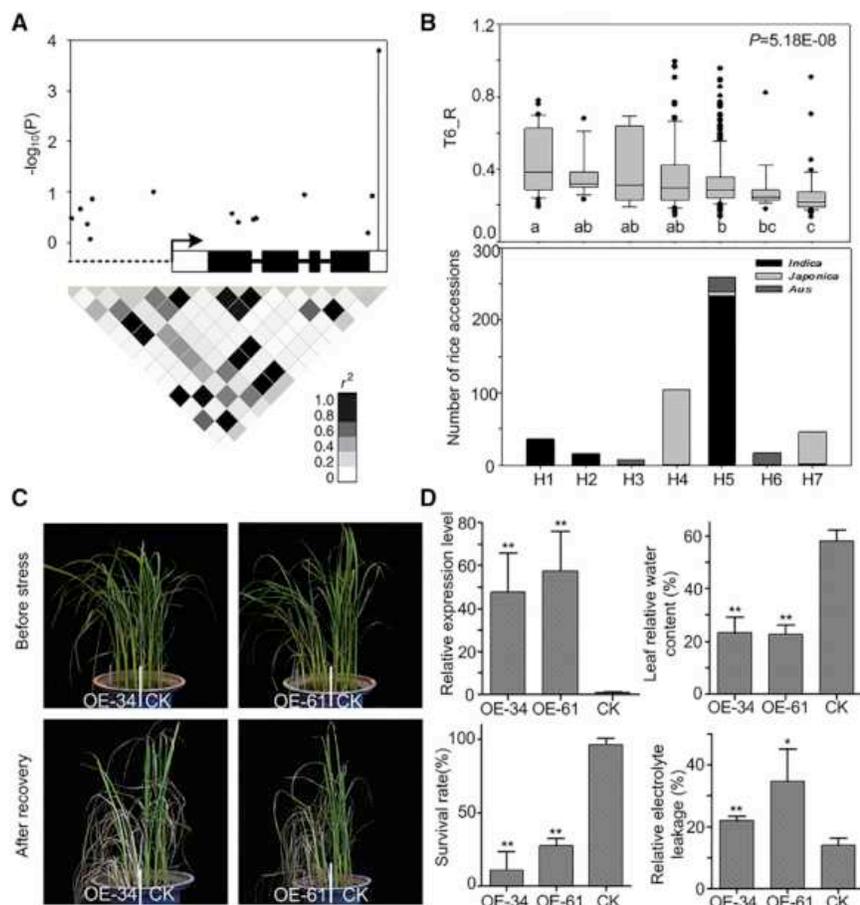


Figure 6. Identification of the New DR-Related Gene *OsPP15*.

(A) Local Manhattan plot and LD statistic r^2 for *OsPP15* (gene body and 1 kb upstream of the gene). The dotted line, solid lines, white rectangles, and black rectangles represent the 1-kb promoter region, introns, UTRs, and exons, respectively. The arrow indicates the transcription start site and transcription orientation.

(B) Comparison of T6_R between the seven haplotype groups of *OsPP15*. The haplotypes were determined based on *OsPP15* SNPs with a MAF ≥ 0.05 . The distribution of T6_R for each haplotype group is shown in the box plot. The i-traits of the seven haplotype groups were compared using Kruskal–Wallis one-way ANOVA. The number of rice accessions in each haplotype group is shown on the histogram.

(C) Images of two independent *OsPP15*-over-expression lines (OE-34 and OE-61) and a negative transgenic line (control) before drought stress and after rewatering. The OE lines were more sensitive to drought stress than the negative control.

(D) Relative expression level, relative leaf water content, and relative electrolyte leakage under drought stress, and the survival rate after rewatering are shown for three lines (OE-34, OE-61, and negative transgenic control). Error bars indicate the SD of three biological replicates and asterisks indicate significance (** $P < 0.01$, * $P < 0.05$; *t*-test).

GPAR_R and relative yield per plant was 0.80 and 0.34, respectively, and GPAR_R was significantly correlated with relative yield per plant in both years ($R = 0.36$, $P < 0.001$ in 2013; $R = 0.42$, $P < 0.001$ in 2016). Although the population size was smaller in 2016 and environmental conditions were different across the two years, 66.6% (313/470) of the loci identified in 2013 were also identified in 2016 due to the high heritability of the i-traits (Supplemental Table 10). For example, SNP sf0801828823 was associated with GPAR_R in both years ($P_{LMM} = 2.88 \times 10^{-12}$ in 2013; $P_{LMM} = 2.06 \times 10^{-5}$ in 2016), and the genotype T allele of the lead SNP was the superior allele for GPAR_R in both years (Figure 7A).

As described above, most DR-related loci co-localized with previously reported DR-related QTLs. To further confirm the usefulness of the i-traits in genetic studies, we conducted drought phenotyping of a biparental mapping population with 192 RILs using the RAP platform and performed QTL mapping analyses. Several QTLs with high significance levels were identified using i-traits_R. For example, a strong QTL on chromosome 1 (peak at 84.01 cM, logarithm of odds [LOD] = 5.69) controlling GPAR_R and a strong QTL on chromosome 2 (peak at 94.81 cM, LOD = 7.31) controlling TBR_R were identified (Supplemental Figure 2). Strikingly, we found that some QTLs identified by linkage analysis overlapped with GWAS loci. For example, a QTL on chromosome 3 (identified by linkage analysis) controlling both the i-trait GPA and green leaf area (measured using HLS in a destructive manner) under drought conditions

co-localized with a locus (identified by GWAS) significantly associated with both traits. GWAS had obviously higher mapping resolution than the linkage analysis (Figure 7B). A total of 69 i-trait locus associations were identified by both GWAS and biparental QTL mapping (Supplemental Table 11).

Relevance of I-Traits to DR Performance in the Field

We further examined whether the i-traits are relevant to DR performance in the field. The same association population was phenotyped for DR in the field, facilitated by a movable rain-off shelter, in 2011 and 2016. Relative yield (stress/non-stress) and two leaf-rolling traits were measured. The two leaf-rolling traits included the number of days from the start of drought treatment to the start of leaf rolling (“days to leaf rolling” for short) and the number of days from the start of leaf rolling to irreversible leaf rolling in the morning (“days during leaf rolling” for short). Drought stress was initiated at the early panicle development stage (see Methods). Strikingly, we found many significant phenotypic correlations between the i-traits and the traditional DR-related traits in the field. For example, GPAR_R and GPA_R were significantly correlated with relative yield, and TBR_R and PC6_R were significantly correlated with the leaf-rolling traits (Figure 7C and Supplemental Table 12). A total of 95 loci significantly associated with i-traits_R were also associated with relative yield and leaf-rolling traits in the field ($P_{LMM} < 1 \times 10^{-3}$), and the i-traits_R had higher power to detect these loci than traditional DR traits (Supplemental Table 13). For example, the lead

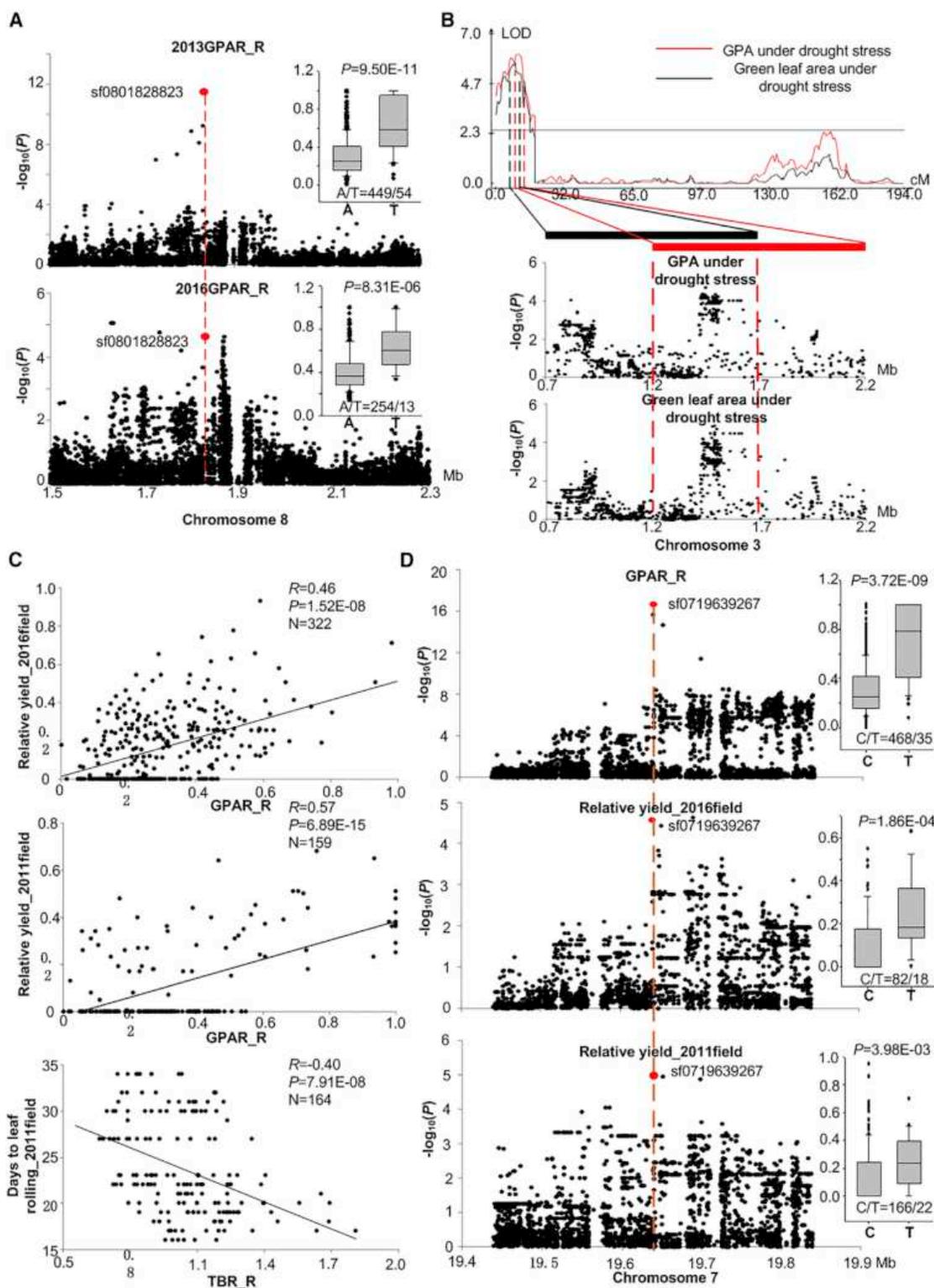


Figure 7. Validation by Repeated GWAS, Linkage Analyses of I-Traits, and DR Phenotyping in the Field.

(A) Local Manhattan plots of the same genomic region associated with GPAR_R across 2 years (2013 and 2016) and box plots showing the GPAR_R distribution for two genotype groups with different alleles of SNP sf0801828823. The GPAR_R values of the two genotype groups were compared using a *t*-test.

(B) Linkage analyses and GWAS of GPA (a stay-green-related i-trait) and green leaf area (measured by HLS in a destructive fashion) under drought stress. The QTL on chromosome 3 controlling both traits (identified by linkage analyses) co-localized with the locus associated with these traits identified by GWAS.

(legend continued on next page)

Molecular Plant

SNP sf0719639267 was associated with GPAR_R ($P_{LMM} = 2.01 \times 10^{-17}$) and relative yield in the field in both years ($P_{LMM} = 1.10 \times 10^{-5}$ in 2011, $P_{LMM} = 2.67 \times 10^{-5}$ in 2016), and the genotype T SNP allele was the superior allele for both traits (Figure 7D). These results suggest that i-traits are also useful for quantifying DR in the field.

DISCUSSION

DR has been regarded as a very complex trait involving dynamic and diverse responses that are controlled by a large number of small-effect loci. Small-effect loci can be difficult to detect using traditional DR-related traits. Several factors, such as heading date variation and heterogeneity of soil water content, severely affect the accuracy of DR phenotyping. New DR-related traits with high heritability that are correlated with yield performance under drought conditions in the field are urgently needed (Hu and Xiong, 2014). Recently, a time-series of transcriptional, physiological, and metabolic data were acquired from *Arabidopsis* plants during a slow transition from well-watered to drought conditions and used to dissect early responses to drought (Bechtold et al., 2016). Several other studies have investigated the genetic basis of DT using different traits. For example, GWAS of drought-induced proline accumulation in *Arabidopsis* was performed and led to the identification of a new proline effector gene (Verslues et al., 2014). GWAS of ABA levels under drought conditions has also been performed in *Arabidopsis* to understand the drought-sensing and signaling mechanism (Kalladan et al., 2017). Because optics-based phenotyping techniques facilitate the measurement of not only traditional traits but also new digital traits (Yang et al., 2014), these techniques provide opportunities to identify new heritable DR-related traits. In this study, we identified 51 drought-induced i-traits with high heritability. These i-traits can be used to monitor dynamic responses to drought and evaluate DR levels. Importantly, our results demonstrate that i-traits are useful for predicting DR performance in the field.

In our study, hundreds of DR-related loci were identified by GWAS of i-traits, including some *a priori* DR-related genes. Most of these loci co-localized with previously reported DR-related QTLs and were identified again in a second year due to the high heritability of i-traits. An association network defined using GWAS results demonstrated the necessity and practicability of dissecting complex DR into heritable and simple i-traits. In this study, we focused on two hub i-traits (GPAR_R and TBR_R), which reflect stay-green and leaf rolling, respectively. Since stay-green and leaf rolling are major indices reflecting DT and DA, respectively, GPAR_R and TBR_R may be very useful for evaluating DT and DA in a high-throughput manner. For GPAR_R, most of the associated loci overlapped with previously reported QTL controlling DT-related traits including leaf drying, osmotic adjustment (Blum, 2017), and cell membrane stability. For TBR_R, most of the associated loci overlapped with QTL controlling DA-related traits including leaf-rolling score, canopy

Genetic Architecture of Drought Resistance in Rice

temperature, and leaf relative water content. These results strongly support the usefulness of these two hub i-traits (GPAR_R and TBR_R) in the evaluation of DT and DA, respectively.

Abundant diversity in rice facilitates the identification of superior alleles for breeding. In addition to GWAS, i-traits were used for biparental QTL mapping and several strong DR-related QTLs were identified. To further examine the relevance of i-traits to DR performance in the field, we performed drought phenotyping of the same population in the field across two years. By comparing the phenotypic data and GWAS results between i-traits_R and traditional DR traits in the field, we found that some i-traits were significantly correlated with DR performance in the field. Many loci identified by mapping i-traits were associated with traditional traits in the field (such as relative yield). These results suggest that the i-traits could be useful for DR prediction when we have difficulty in phenotyping a large number of genetic mapping or breeding populations in the field without appropriate controls for environmental factors. Our study also provides a new approach (i-trait-based GWAS) to reveal the genetic architecture of complex traits.

In comparison with biparental QTL mapping, GWAS can provide higher mapping resolution (Huang and Han, 2014). Despite the high mapping resolution, it is difficult to reveal a causal gene underlying the association signal due to slow LD decay in rice. In this study, using the gene network data from RiceNet and GWAS results, an unreported DR-related gene, *OsP15*, was identified, and its function in DR was confirmed by haplotype analyses and genetic transformation experiments. This work provides an example of fast identification and verification of candidate genes for DR through i-trait-based genetic analysis.

METHODS

Plant Material and Experimental Design

A total of 529 *O. sativa* accessions including landraces and elite varieties were used in this study. Of these accessions, 22 accessions that were extremely small or imposed of inaccurate stress treatment, were excluded for subsequent data analyses. Detailed information about the accessions and genotypes has been reported in previous studies (Xie et al., 2015; Zhao et al., 2015). The remaining 507 accessions included 292 *indica* accessions and 139 *japonica* accessions that were classified into two groups based on heading date. The germination dates were staggered to allow these groups to flower synchronously. Seeds were sown in the field and 20-day-old seedlings were transplanted to the greenhouse of a high-throughput rice phenotyping facility at Huazhong Agricultural University in 2013. In 2016, 300 diverse accessions, with abundant genetic diversity and large phenotypic variation of DR (based on the results from 2013), were planted for a second round of drought phenotyping.

Phenotyping

The phenotyping facility used in the study contained the following three sections: RAP, HLS, and YTS. The rice plants were transported to the RAP, which captured 14 color images from different angles, then 51

(C) Scatter diagram showing the correlation between GPAR_R and relative yield (stress/non-stress) in the field in two different years (2011 and 2016), and between TBR_R and days to leaf rolling (number of days from the start of drought treatment to the start of leaf rolling) in the field in 2011.

(D) Local Manhattan plots of the same genomic region associated with GPAR_R and relative yield in the field in two different years and box plots showing the distribution of these traits for two genotype groups with different alleles of SNP sf0719639267. The trait values of the two genotype groups were compared using a *t*-test.

image-based traits (termed “i-traits,” including one biomass-related, four color-related, 25 morphological-related, and 21 texture-related features) were extracted through image analyses. After the rice leaves were cut and fed to the HLS, 10 leaf color-related traits were obtained. After harvesting seeds, 11 yield traits were measured by YTS. Biomass-related traits were measured by weighing.

To determine whether i-traits measured by RAP are useful measures of the dynamic changes of rice plants in response to drought stress, we phenotyped three plants of a rice variety (Swarna) at eight time points during progressive drought stress and rewatering. The soil water content was monitored by time domain reflectometry (TDR) using a TRIME-PICO32 (IMKO Micromodultechnik, Ettlingen, Germany). After the first round of phenotyping, irrigation was stopped immediately; therefore, the phenotypes measured at day 0 corresponded to the phenotypes before drought stress. Moderate stress (the first round of drought stress applied for 2 days) and severe stress (the second round of drought stress applied for 3 days) were applied. The soil water content during the second round of stress was lower than that during the first round of stress. Based on images of rice plants, chlorophyll began to degrade during the second round of stress but not during the first round of stress.

For phenotyping the natural population (507 accessions were actually used), a total of eight healthy plants for each accession were grown in pots filled with 4.5 kg of soil per pot. Four plants for each accession were subjected to drought stress at the panicle development stage, at which rice is most sensitive to drought stress. The experiment followed a randomized complete block design and each of the four blocks contained one plant of each accession. Two plants of each accession were used to trace panicle development in a destructive way. The remaining two plants were used to quantify the yield traits under normal conditions to obtain the relative yield-related traits (the ratios of yield-related trait values under drought stress to those under normal conditions), which can better reflect DR. However, unexpectedly high temperatures from July to August at Wuhan resulted in low spikelet fertility under normal conditions (Supplemental Figure 3). To reduce the effect of high temperatures on yield traits, we excluded rice accessions with spikelet fertility <60% under normal conditions when calculating the relative yield traits. Therefore, the relative yield-related traits for 222 accessions were retained for subsequent analyses.

When rice plants grew to the booting stage (panicle elongation), the four plants designated for stress treatment were phenotyped by RAP before stress treatment. Irrigation was then stopped to impose drought stress. The soil water content was monitored by TDR using a TRIME-PICO32 (IMKO Micromodultechnik). When the soil water content decreased to 15% (TDR value), the plants were watered once per day to maintain the soil water content at 15% (TDR value) for 5 days. Four replicates of each accession were then phenotyped by RAP again to collect phenotype data under drought stress. The interval between the two rounds of phenotyping was approximately 1 week, depending on the rate of decline in soil water content. After the second round of phenotyping by RAP, two replicates were immediately sampled for measurements of leaf color-related traits by HLS and shoot weight-related traits by weighing. The remaining two replicates were rewatered at the same time. When the accessions entered the maturity stage, the two replicates were phenotyped by RAP and then harvested for measurements of yield traits by YTS and biomass-related traits by weighing. Additionally, two replicates grown under normal conditions were harvested and yield traits and biomass were measured to calculate relative yield traits and relative biomass (the ratios of traits under drought/normal conditions). Drought phenotyping of 300 rice accessions in 2016 was performed in the same way as in 2013.

Image Acquisition and Processing

Fourteen side-view images from different angles for each rice plant were acquired using a charge-coupled device camera (Stingray F-504C;

Applied Vision Technologies, Germany). The image processing was performed using LabVIEW (National Instruments, USA).

Image Segmentation

The original RGB image was transformed to HSI color space. Background pixels and plant pixels were discriminated using fixed thresholds. A binary image of the plant was generated by setting plant pixels as one and background pixels as 0.

Color Component Extraction

When the binary image of the plant was used as a mask, an RGB image without background was generated from the original RGB image. The ExG (excessive green) and ExR (excessive red) planes of the RGB image were extracted to determine the green part of the rice plant.

If the ExG value was greater than a predefined ExG threshold and the ExR value was less than a predefined ExR threshold, the corresponding pixels were defined as greenness pixels. The i2 component was extracted to determine yellow part of the rice plant.

Edge Detection and Bounding Rectangle Detection

The edge of the plant was extracted using IMAQ EdgeDetection VI. The bounding rectangle was defined as the rectangle of minimum area that surrounded a plant, and was detected using IMAQ Particle Analysis VI.

The equations are shown in Supplemental Table 1.

I-Trait Calculations

- Total projected area (TPA): Number of foreground pixels attributed to the rice plant.
- Green projected area (GPA): Number of foreground pixels attributed to the green part of the rice plant.
- Green projected area ratio (GPAR): GPA/TPA.
- Light-green projected area (LGPA): Number of foreground pixels attributed to the pale-green part of the rice plant. The HLS segmentation values of the pale-green part in an image were identified using a standard color chart.
- Dark-green projected area (DGPA): Number of foreground pixels attributed to the dark-green part of the rice plant. The HLS segmentation values of the dark-green part in an image were also identified using a standard color chart.
- Fractal dimension without image cropping (FDNIC): Boxes with a box size of δ_k were superimposed on the object of interest, and the number of boxes that were needed to cover the object, denoted as N_{δ_k} was calculated. This process was repeated with decreasing δ_k until δ_k approached pixel size. The equation is shown in Supplemental Table 1.
- Fractal dimension after image cropping (FDIC): The original image was cropped to a smaller size with the bounding rectangle of the rice plant, and the FD was calculated using the above steps.
- H: Height of the bounding rectangle of the object.
- W: Width of the bounding rectangle of the object.
- Height/width ratio (HWR): H/W.
- Total projected area/bounding rectangle area ratio (TBR): $A/(H \times W)$
- Perimeter/projected area ratio (PAR): The length of the outline of a rice plant was calculated, then the length was divided by the projected area.
- Plant compactness (PC1–PC6): The image was divided into several subimages using a (5 × 5) window. The ratio of the number of foreground pixels to the total number of pixels in each subimage (5 × 5) was calculated and denoted as plant compactness in each subimage (PCs). PCs were categorized into six classes: C1: <10%, C2: 10%–20%, C3: 20%–40%, C4: 40%–60%, C5: 60%–80%, C6: 80%–100%. Then the number of PCs belonging to each class, denoted as ND_i ($i = 1, 2, \dots, 6$), was counted. Lastly, leaf compactness of class i (PCi) was computed as the percentage of ND_i compared with the sum of ND_i . The levels of PC3 under drought stress were not significantly different from the levels under normal conditions

Molecular Plant

($P > 0.05$, paired-samples t -test), so this trait was not included in the study.

- Relative frequencies (F1–F14):
The equations are shown in [Supplemental Table 1](#).
The levels of F11 under drought stress were not significantly different from the levels under normal conditions ($P > 0.05$, paired-samples t -test), so this trait was not included in the study.
- Six histogram traits: the mean value (M), the SE, the third moment (MU3), the uniformity (U), the smoothness (S), and the entropy (E). The equations are shown in [Supplemental Table 1](#).
- Fifteen gray-level co-occurrence matrix texture traits: the correlation (T1), the advantages of the small gradient (T2), the advantages of the large gradient (T3), the energy (T4), the intensity inhomogeneity (T5), the gradient inhomogeneity (T6), the mean gray (T7), the mean gradient (T8), the gray entropy (T9), the gradient entropy (T10), the entropy of mixing (T11), the differential moment (T12), the deficit score (T13), the gray variance (T14), and the gradient variance (T15). The equations are shown in [Supplemental Table 1](#).

Leaf Traits Measured by HLS

- Green leaf area (GLA): Represents the area of green leaves per plant.
- Total leaf area (TLA): Represents the leaf area (including green leaves and yellow leaves) per plant.
- Green leaf area ratio: GLA/TLA.
- Yellow leaf area: Represents the total area of yellow leaves per plant.
- Green-2 leaf area (GLA₂): Represents the area of leaves whose color is similar to group 2 of the leaf color chart designed by the IRR1.
- Green-2 leaf area ratio: GLA₂/GLA.
- Green-3 leaf area (GLA₃): Represents the area of leaves whose color is similar to group 3 of the leaf color chart designed by the IRR1.
- Green-3 leaf area ratio: GLA₃/GLA.
- Green-4 leaf area (GLA₄): Represents the area of leaves whose color is similar to group 4 of the leaf color chart designed by the IRR1.
- Green-4 leaf area ratio: GLA₄/GLA.

Analyses of Phenotypic Data

The broad-sense heritability (H^2) was calculated using ANOVA based on the phenotypic data from 2 years: $H^2 = V_G/(V_G + V_e/N)$; phenotypic variance was partitioned into genotypic (V_G) and environmental (V_e) variance ([Chen et al., 2014b](#)); N represents the number of years ($N = 2$ in this study). The repeatability (w^2) was calculated using ANOVA from the phenotypic data for replicated plants in 2013: $w^2 = V_g/(V_g + V_e/n)$; phenotypic variance was partitioned into genotypic (V_g) variance and variance due to differences in repeated measures (V_e); n represents the number of replicated plants of the same accession ($n = 4, 4, 2$ for traits measured before drought stress, under drought stress, and after rewatering, respectively) ([Chen et al., 2014a](#)).

Paired-sample and independent-sample t -tests, stepwise linear regression, and calculation of pairwise Pearson correlation coefficients (R) were performed using IBM SPSS version 19 (IBM, Armonk, USA).

A heatmap was drawn with Heml software ([Deng et al., 2014](#)). Before the heatmap was drawn, the raw phenotypic data were linearly normalized as $y = (x - \min)/(\max - \min)$ in which x , y , \max , and \min represent raw data, normalized data, maximum, and minimum, respectively.

All-subset regression was adopted for the modeling of shoot weight, and adjusted R^2 (adjusted determination coefficients), AIC (Akaike's information criterion), MAPE (mean absolute percentage error), and SD_{APE} (the SD of absolute percentage error) were calculated using SAS 9.2 (SAS Institute). To evaluate the prediction performance of a model, we independently performed 5-fold cross-validation ([Hallmark et al., 2007](#); [Yang et al.,](#)

Genetic Architecture of Drought Resistance in Rice

2014) 10 times. For 5-fold cross-validation, five steps were conducted as follows: (i) the rice accessions were randomly assigned into five groups; (ii) one of the five groups was selected as a testing set and the other four groups were selected as a training set; (iii) MAPE, SD_{APE} , and R^2 of the testing set were calculated; (iv) the final MAPE, SD_{APE} , and R^2 were calculated as the mean values of 10 reruns of 5-fold cross-validation.

Genome-Wide Association Study

A total of 4 358 600, 2 863 169, and 1 959 460 SNPs (minor allele frequency ≥ 0.05 ; the number of accessions with minor alleles ≥ 6) were used for GWAS of the whole population, the *indica* subpopulation, and the *japonica* subpopulation, respectively. The genotypes of these SNPs were obtained from the RiceVarMap database (<http://ricevarmap.ncpgr.cn/>) ([Zhao et al., 2015](#)). The mean calculated from replicates for each trait was used for the GWAS. The same trait measured at different stages was regarded as a different trait for association analyses. Ratio traits were calculated by dividing the trait value under stress by the value under non-stress conditions. In total, we performed GWAS for 255 traits, including i-traits measured by RAP and traditional traits measured by HLS, YTS, and manual measurement. To better reflect drought responses or DR, we mainly focused on the ratio traits in this study.

Considering the small effects of DR-related loci and the low heritability of DR, the suggestive P value threshold ($1/n$) was set to control the genome-wide type I error rate of GWAS ([Duggal et al., 2008](#); [Li et al., 2013](#); [Yang et al., 2014](#); [Wang et al., 2015](#)); n represents the effective number of independent SNPs calculated by the GEC software tool ([Li et al., 2012](#)). The thresholds were 1.21×10^{-6} , 1.66×10^{-6} , and 3.81×10^{-6} for the whole population, the *indica* subpopulation, and the *japonica* subpopulation, respectively. GWAS was performed using a mixed-model approach with the factored spectrally transformed linear mixed models (FaST-LMM) program ([Lippert et al., 2011](#)), and independent lead SNPs were obtained by using the “clumping” function of Plink ([Purcell et al., 2007](#)). All genes within 100 kb both upstream and downstream of the lead SNPs were extracted to identify potential candidate genes. Considering the slow LD decay in rice and based on previously published studies ([Chen et al., 2014b](#); [Yang et al., 2015](#); [Crowell et al., 2016](#)), adjacent lead SNPs in a region less than 300 kb were defined as a single locus. An association network based on associations between loci and DR-related traits (mainly composed of ratio traits) was displayed using Cytoscape software (<http://www.cytoscape.org/>) ([Shannon et al., 2003](#)). To test whether the DR-related loci were randomly distributed in the genome, we partitioned the whole genome into 2-Mb segments, counted the number of loci in each segment, and performed a χ^2 test using SPSS version 19 (IBM).

Haplotypes were determined based on the genotypes of SNPs with a MAF ≥ 0.05 , and the LD statistic r^2 was calculated using Haploview 4.2 ([Barrett et al., 2005](#)). Multiple comparisons of groups of accessions with different haplotypes were conducted using a non-parametric test (Kruskal–Wallis one-way ANOVA) using SPSS version 19 (IBM).

Retrieval of Reported DR-Related QTLs from Databases

Previously reported DR-related QTLs were retrieved from TropGeneDB (<http://TropGeneDBdb.cirad.fr/TropGeneDB/JSP/interface.jsp?module=RICE>), QTARO (<http://qtaro.abr.affrc.go.jp/>), and PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>). A total of 387 QTLs were retrieved by selecting “DROUGHT: non root traits” (“Trait” menu) and “Water stress” (“Water condition” menu) in TropGeneDB; a total of 111 QTLs were retrieved by selecting “Resistance or Tolerance” (“Major category” menu) and “Drought tolerance” (“Category of object character” menu) in QTARO; a total of 239 QTLs were retrieved from the literature by inputting “drought” “QTL” “rice” (“Text Word” menu) in PubMed.

Identification of New DR-Related Genes

The web server RiceNet was used to predict new candidate genes for a phenotype or pathway by submitting guide genes with known functions (<http://www.inetbio.org/ricenet>). A total of 21 different datasets are integrated into RiceNet (v 2.0). These datasets contain information about co-expression, protein-protein interactions, and genetic interactions in diverse species. The genes that are closely connected to the guide genes were regarded as potential new candidates. In this study two known DR-related genes, *OsNCED3* (LOC_Os03g44380) and *OsDREB1E* (LOC_Os04g48350), were located in loci associated with T6_R.

In this study, the unreported gene *OsPP15* (LOC_Os01g62760) was predicted to be connected with the two known DR-related genes based on RiceNet. Since *OsPP15* was associated with T6_R based on GWAS, this gene was selected as a new DR-related candidate gene for further genetic transformation experiments.

Genetic Transformation and Functional Validation of *OsPP15*

A DNA fragment harboring the coding region of the candidate gene *OsPP15* was amplified from the genomic DNA of *japonica* rice cv. Zhonghua 11 using primers PP15-1301F (5'-TAC GAA CGA TAG CCG GTA CCA TGG CCG AGA TCT GCT GCG AG-3') and PP15-1301R (5'-TTG CGG ACT CTA GAG GAT CCT CAC AAT CCC CGG CGG AGA TC-3'), and the product was cloned into the binary vector pCAMBIA1301U (digested by *KpnI* and *BamHI*) by Gibson assembly (Gibson et al., 2009). The sequence-confirmed construct was transformed into Zhonghua 11 via *Agrobacterium*-mediated transformation according to standard methods (Hiei et al., 1994). To identify overexpression and negative transgenic lines, we quantified the expression level of *OsPP15* at the T₀ generation by real-time qRT-PCR using primers PP15-qF2 (5'-TAT GGG ACG TCG TGA CCA AC-3') and PP15-qR2 (5'-AGG ACT CCG CCT TGC TTA TC-3').

Two independent overexpression lines (OE-34 and OE-61) and a negative transgenic line (CK) were planted in pots and treated with drought stress at the 4-leaf stage. When the leaves of CK rolled due to drought stress (soil water content of 13.9% ± 1.6% [water (g)/dry soil (g)]), the stressed seedlings of the three lines (OE-34, OE-61, CK) were sampled for the measurement of leaf relative water content (RWC) and relative electrolyte leakage (REL) (three replicates for each line). The stressed seedlings in other pots were rewatered and the survival rates were calculated (three replicates for each line).

Leaf RWC was measured using the following procedure: (i) the leaves were sampled and weighed immediately to obtain the fresh weight (FW); (ii) the leaves were immersed in distilled water in darkness for 24 h to obtain the saturated weight (SW); (iii) the leaves were dried in an oven at 80°C for 48 h to obtain the dry weight (DW); (iv) the leaf RWC was calculated: $RWC (\%) = (FW - DW) / (SW - DW) \times 100$.

Leaf REL was measured using the following procedure: (i) the leaf samples were cut into small pieces and immersed in tubes filled with distilled water at 25°C for 24 h; (ii) the conductivity (RL1) was measured using a conductivity meter (Model DDSIIA, Shanghai Leici Instrument, Shanghai, China); (iii) the tubes were placed in boiling water for 20 min and cooled naturally to room temperature; (iv) the conductivity (EL2) was measured; (v) REL was calculated: $REL (\%) = EL1/EL2 \times 100$.

Biparental QTL Mapping

The QTL mapping population containing 192 RILs was derived from a cross between the upland rice variety IRAT 109 and the lowland rice variety Zhenshan 97 (Zou et al., 2005). The genotypic data for this population were provided by Shanghai Agrobiological Gene Center. The linkage map is 1567 centimorgans in length and contains 2499 bins, with 0.63 cM per bin on average. The RIL population, four plants per line, was phenotyped for DR at the high-throughput phenotyping facility at Huazhong Agricultural University in 2016. The drought treatment and phenotyping proced-

ure were the same as for association mapping as described above. QTL mapping was performed using the composite interval mapping method (Model 6: Standard Model) with WinQTLCart v2.5 with a 0.5 cM walking speed (Wang et al., 2010). The backward regression method was selected. The control marker number and window size were set to 5 and 10 cM, respectively, for background controls. The LOD threshold was set to 2.5 and a two-LOD drop support interval was used for each QTL.

Drought Phenotyping in the Field

Due to the large variation in heading date among rice accessions phenotyped with RAP (a big obstacle for accurate drought evaluation at the reproductive stage in the field) and the limited area of the facility field site, only 197 and 330 accessions were planted in 2011 and 2016, respectively, for drought phenotyping in the field under a movable rain-off shelter. The experiments followed a randomized complete block design with two replications and two treatments (drought and normal growth). Field cultivation management was done following the standard protocol for middle season rice production in central China. For each accession, 20 plants were grown for each of the two replications. Drought stress was initiated at the early stage of panicle development (2–3 mm in length) by stopping irrigation and closing the shelter if raining. The field was then irrigated when 70% of the accessions showed irreversible leaf-rolling observed in the morning. During progressive drought stress, the number of days from the start of drought treatment to the start of leaf rolling (“days to leaf rolling” for short) and the number of days from the start of leaf rolling to irreversible leaf rolling in the morning (“days during leaf rolling” for short) were recorded for each accession in 2011. The grains in each plot were harvested and weighed to calculate the relative yield (stress/non-stress) in 2011 and 2016. GWAS of the traits in the field was performed as described above.

Phenomics Database

All images, image processing algorithms for i-traits, and phenotypic data can be downloaded from our phenomics database: http://plantphenomics.hzau.edu.cn/search_en.action.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

We thank the National Key Research and Development Program of China (2016YFD0100600, 2016YFD0100101-18), the National Program on High Technology Development (2014AA10A600), the National Natural Science Foundation of China (31770397), and the Fundamental Research Funds for the Central Universities (2662015PY126, 2662016PY0092, 662017PY058). We also thank Prof. John Doonan of Aberystwyth University for the English language improvement.

AUTHOR CONTRIBUTIONS

Z.G. designed and performed the experiments, conducted GWAS and data analysis, and wrote the manuscript; W.Y. extracted i-traits from raw images; Y.C. generated transgenic rice overexpression lines; F.X., N.J., H.F., C.H., and P.Y. participated in the phenotyping; X.M., H.L., and L.L. provided and genotyped the RIL population; H.T. provided the yield data for 2016; H.Z. provided genotypic data; G.C. contributed to the rice cultivation; H.H. revised the manuscript; Q.L. participated in designing the RAP platform; L.X. conceived the project, supervised the study, and wrote the manuscript. All of the authors carefully read and discussed the manuscript.

ACKNOWLEDGMENTS

No conflict of interest declared.

Received: October 30, 2017
Revised: February 14, 2018
Accepted: March 26, 2018
Published: March 30, 2018

REFERENCES

- Alam, M.M., Mace, E.S., van Oosterom, E.J., Cruickshank, A., Hunt, C.H., Hammer, G.L., and Jordan, D.R. (2014). QTL analysis in multiple sorghum populations facilitates the dissection of the genetic and physiological control of tillering. *Theor. Appl. Genet.* **127**:2253–2266.
- Bac-Molenaar, J.A., Vreugdenhil, D., Granier, C., and Keurentjes, J.J. (2015). Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci. *J. Exp. Bot.* **66**:5567–5580.
- Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**:263–265.
- Bechtold, U., Penfold, C.A., Jenkins, D.J., Legaie, R., Moore, J.D., Lawson, T., Matthews, J.S.A., Violet-Chabrand, S.R.M., Baxter, L., Subramaniam, S., et al. (2016). Time-series transcriptomics reveals that AGAMOUS-LIKE22 affects primary metabolism and developmental processes in drought-stressed *Arabidopsis*. *Plant Cell* **28**:345–366.
- Berger, B., Parent, B., and Tester, M. (2010). High-throughput shoot imaging to study drought responses. *J. Exp. Bot.* **61**:3519–3528.
- Blum, A. (2011). Drought resistance—is it really a complex trait? *Funct. Plant Biol.* **38**:753–757.
- Blum, A. (2017). Osmotic adjustment is a prime drought stress adaptive engine in support of plant production. *Plant Cell Environ.* **40**:4–10.
- Busemeyer, L., Ruckelshausen, A., Moller, K., Melchinger, A.E., Alheit, K.V., Maurer, H.P., Hahn, V., Weissmann, E.A., Reif, J.C., and Wurschum, T. (2013). Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. *Sci. Rep.* **3**:2442.
- Cai, S., Jiang, G., Ye, N., Chu, Z., Xu, X., Zhang, J., and Zhu, G. (2015). A key ABA catabolic gene, OsABA8ox3, is involved in drought stress resistance in rice. *PLoS One* **10**:e0116646.
- Campbell, M.T., Du, Q., Liu, K., Brien, C.J., Berger, B., Zhang, C., and Walia, H. (2017). A comprehensive image-based phenomic analysis reveals the complex genetic architecture of shoot growth dynamics in rice (*Oryza sativa*). *Plant Genome* **10**. <https://doi.org/10.3835/plantgenome2016.07.0064>.
- Chen, J.Q., Meng, X.P., Zhang, Y., Xia, M., and Wang, X.P. (2008). Over-expression of OsDREB genes lead to enhanced drought tolerance in rice. *Biotechnol. Lett.* **30**:2191–2198.
- Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., and Klukas, C. (2014a). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* **26**:4636–4655.
- Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., Li, Y., Liu, X., Zhang, H., Dong, H., et al. (2014b). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* **46**:714–721.
- Crasta, O.R., Xu, W.W., Rosenow, D.T., Mullet, J., and Nguyen, H.T. (1999). Mapping of post-flowering drought resistance traits in grain sorghum: association between QTLs influencing premature senescence and maturity. *Mol. Gen. Genet.* **262**:579–588.
- Crowell, S., Korniliev, P., Falcao, A., Ismail, A., Gregorio, G., Mezey, J., and McCouch, S. (2016). Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nat. Commun.* **7**:10527.
- Cui, M., Zhang, W., Zhang, Q., Xu, Z., Zhu, Z., Duan, F., and Wu, R. (2011). Induced over-expression of the transcription factor OsDREB2A improves drought tolerance in rice. *Plant Physiol. Biochem.* **49**:1384–1391.
- Deng, W.K., Wang, Y.B., Liu, Z.X., Cheng, H., and Xue, Y. (2014). Heml: A toolkit for illustrating heatmaps. *PLoS One* **9**:e111988.
- Duggal, P., Gillanders, E.M., Holmes, T.N., and Bailey-Wilson, J.E. (2008). Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics* **9**:516.
- Eliasson, J. (2015). The rising pressure of global water shortages. *Nature* **517**:6.
- Fang, Y., and Xiong, L. (2015). General mechanisms of drought response and their application in drought resistance improvement in plants. *Cell Mol. Life Sci.* **72**:673–689.
- Fukao, T., and Xiong, L. (2013). Genetic mechanisms conferring adaptation to submergence and drought in rice: simple or complex? *Curr. Opin. Plant Biol.* **16**:196–204.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**:343–345.
- Hallmark, S.L., Souleyrette, R., and Lamprey, S. (2007). Use of n-fold cross-validation to evaluate three methods to calculate heavy truck annual average daily traffic and vehicle miles traveled. *J. Air Waste Manag. Assoc.* **57**:4–13.
- Hamelin, C., Sempere, G., Jouffe, V., and Ruiz, M. (2013). TropGeneDB, the multi-tropical crop information system updated and extended. *Nucleic Acids Res.* **41**:D1172–D1175.
- Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994). Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**:271–282.
- Hu, H., and Xiong, L. (2014). Genetic engineering and breeding of drought-resistant crops. *Annu. Rev. Plant Biol.* **65**:715–741.
- Huang, X., and Han, B. (2014). Natural variations and genome-wide association studies in crop plants. *Annu. Rev. Plant Biol.* **65**:531–551.
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* **42**:961–967.
- Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., et al. (2011). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* **44**:32–39.
- Johnson, S.M., Cummins, I., Lim, F.L., Slabas, A.R., and Knight, M.R. (2015). Transcriptomic analysis comparing stay-green and senescent *Sorghum bicolor* lines identifies a role for proline biosynthesis in the stay-green trait. *J. Exp. Bot.* **66**:7061–7073.
- Kalladan, R., Lasky, J.R., Chang, T.Z., Sharma, S., Juenger, T.E., and Verslues, P.E. (2017). Natural variation identifies genes affecting drought-induced abscisic acid accumulation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **114**:11536–11541.
- Lanceras, J.C., Pantuwan, G., Jongdee, B., and Toojinda, T. (2004). Quantitative trait loci associated with drought tolerance at reproductive stage in rice. *Plant Physiol.* **135**:384–399.
- Lee, T., Oh, T., Yang, S., Shin, J., Hwang, S., Kim, C.Y., Kim, H., Shim, H., Shim, J.E., Ronald, P.C., et al. (2015). RiceNet v2: an improved network prioritization server for rice genes. *Nucleic Acids Res.* **43**:W122–W127.
- Li, M.X., Yeung, J.M., Cherny, S.S., and Sham, P.C. (2012). Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum. Genet.* **131**:747–756.
- Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., et al. (2013). Genome-wide association study

- dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* **45**:43–50.
- Lippert, C., Listgarten, J., Liu, Y., Kadie, C.M., Davidson, R.I., and Heckerman, D.** (2011). FaST linear mixed models for genome-wide association studies. *Nat. Methods* **8**:833–835.
- Meyer, K., Leube, M.P., and Grill, E.** (1994). A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. *Science* **264**:1452–1455.
- Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., Stein, N., Graner, A., and Kilian, B.** (2015). Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. *Plant Cell Environ.* **38**:1980–1996.
- Price, A.H., Townend, J., Jones, M.P., Audebert, A., and Courtois, B.** (2002). Mapping QTLs associated with drought avoidance in upland rice grown in the Philippines and West Africa. *Plant Mol. Biol.* **48**:683–695.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., et al.** (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**:559–575.
- Riedelsheimer, C., Lisec, J., Czedik-Eysenberg, A., Sulpice, R., Flis, A., Grieder, C., Altmann, T., Stitt, M., Willmitzer, L., and Melchinger, A.E.** (2012). Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. USA* **109**:8872–8877.
- Rosenow, D.T., Quiseberry, J.E., Wendt, C.W., and Clark, L.E.** (1983). Drought tolerant sorghum and cotton germplasm. *Agr. Water Manag.* **7**:207–222.
- Ruiz, M., Rouard, M., Raboin, L.M., Lartaud, M., Lagoda, P., and Courtois, B.** (2004). TropGENE-DB, a multi-tropical crop information system. *Nucleic Acids Res.* **32**:D364–D367.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T.** (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**:2498–2504.
- Tester, M., and Langridge, P.** (2010). Breeding technologies to increase crop production in a changing world. *Science* **327**:818–822.
- Verslues, P.E., Lasky, J.R., Juenger, T.E., Liu, T.W., and Kumar, M.N.** (2014). Genome-wide association mapping combined with reverse genetics identifies new effectors of low water potential-induced proline accumulation in *Arabidopsis*. *Plant Physiol.* **164**:144–159.
- Wang, S., Basten, C.J., and Zeng, Z.** (2010). Windows QTL Cartographer 2.5 (Raleigh, NC: Department of Statistics, North Carolina State University).
- Wang, Q., Xie, W., Xing, H., Yan, J., Meng, X., Li, X., Fu, X., Xu, J., Lian, X., Yu, S., et al.** (2015). Genetic architecture of natural variation in rice chlorophyll content revealed by a genome-wide association study. *Mol. Plant* **8**:946–957.
- Wang, X., Wang, H., Liu, S., Ferjani, A., Li, J., Yan, J., Yang, X., and Qin, F.** (2016). Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings. *Nat. Genet.* **48**:1233–1241.
- Xiao, J., Cheng, H., Li, X., Xiao, J., Xu, C., and Wang, S.** (2013). Rice WRKY13 regulates cross talk between abiotic and biotic stress signaling pathways by selective binding to different cis-elements. *Plant Physiol.* **163**:1868–1882.
- Xie, W., Wang, G., Yuan, M., Yao, W., Lyu, K., Zhao, H., Yang, M., Li, P., Zhang, X., Yuan, J., et al.** (2015). Breeding signatures of rice improvement revealed by a genomic variation map from a large germplasm collection. *Proc. Natl. Acad. Sci. USA* **112**:E5411–E5419.
- Xing, Y., and Zhang, Q.** (2010). Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.* **61**:421–442.
- Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., Fang, W., Feng, H., Xie, W., Lian, X., et al.** (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* **5**:5087.
- Yang, W., Guo, Z., Huang, C., Wang, K., Jiang, N., Feng, H., Chen, G., Liu, Q., and Xiong, L.** (2015). Genome-wide association study of rice (*Oryza sativa* L.) leaf traits with a high-throughput leaf scorer. *J. Exp. Bot.* **66**:5605–5615.
- Yonemaru, J., Yamamoto, T., Fukuoka, S., Uga, Y., Hori, K., and Yano, M.** (2010). Q-TARO: QTL annotation rice online database. *Rice* **3**:194–203.
- Yue, B., Xiong, L., Xue, W., Xing, Y., Luo, L., and Xu, C.** (2005). Genetic analysis for drought resistance of rice at reproductive stage in field with different types of soil. *Theor. Appl. Genet.* **111**:1127–1136.
- Yue, B., Xue, W., Xiong, L., Yu, X., Luo, L., Cui, K., Jin, D., Xing, Y., and Zhang, Q.** (2006). Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. *Genetics* **172**:1213–1228.
- Zhang, Q.** (2007). Strategies for developing green super rice. *Proc. Natl. Acad. Sci. USA* **104**:16402–16409.
- Zhao, K., Tung, C.W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., et al.** (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2**:467.
- Zhao, H., Yao, W., Ouyang, Y., Yang, W., Wang, G., Lian, X., Xing, Y., Chen, L., and Xie, W.** (2015). RiceVarMap: a comprehensive database of rice genomic variations. *Nucleic Acids Res.* **43**:D1018–D1022.
- Zou, G.H., Mei, H.W., Liu, H.Y., Liu, G.L., Hu, S.P., Yu, X.Q., Li, M.S., Wu, J.H., and Luo, L.J.** (2005). Grain yield responses to moisture regimes in a rice population: association among traits and genetic markers. *Theor. Appl. Genet.* **112**:106–113.